Convergence of the Iterative Rational Krylov Algorithm

Garret Flagg, Christopher Beattie, Serkan Gugercin
Department of Mathematics, Virginia Tech.
Blacksburg, VA, 24061-0123
e-mail: {flagg,beattie,gugercin}@math.vt.edu

July 28, 2011

Abstract

Iterative Rational Krylov Algorithm (IRKA) of [8] is an interpolatory model reduction approach to optimal $\mathcal{H}_2$ approximation problem. Even though the method has been illustrated to show rapid convergence in various examples, a proof of convergence has not been provided yet. In this note, we show that in the case of state-space symmetric systems, IRKA is a locally convergent fixed point iteration to a local minimum of the underlying $\mathcal{H}_2$ approximation problem.

1 Introduction

Consider a single-input-single-output (SISO) linear dynamical system in state-space form:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\,\boldsymbol{x}(t) + \boldsymbol{b}\,u(t), \qquad y(t) = \boldsymbol{c}^T\boldsymbol{x}(t), \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, and $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$. In (1), $\boldsymbol{x}(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, are, respectively, the *states*, *input*, and *output* of the dynamical system. The transfer function of the underlying system is $H(s) = \boldsymbol{c}^T(s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b}$. $H(s)$ will be used to denote both the system and its transfer function.

Dynamical systems of the form (1) with large state-space dimension n appear in many applications; see, e.g., [1] and [10]. Simulations in such large-scale settings make enormous demands on computational resources. The goal of model reduction is to construct a surrogate system

$$\dot{\boldsymbol{x}}_r(t) = \boldsymbol{A}_r\,\boldsymbol{x}_r(t) + \boldsymbol{b}_r u(t), \qquad y_r(t) = \boldsymbol{c}_r^T\boldsymbol{x}_r(t), \tag{2}$$

of much smaller dimension $r \ll n$, with $\boldsymbol{A}_r \in \mathbb{R}^{r\times r}$ and $\boldsymbol{b}_r$, $\boldsymbol{c}_r \in \mathbb{R}^r$ such that $y_r(t)$ approximates $y(t)$ well in a certain norm. Similar to $H(s)$, the transfer function $H_r(s)$ of the reduced-model (2) is given by $H_r(s) = \boldsymbol{c}_r^T(s\boldsymbol{I}_r - \boldsymbol{A}_r)^{-1}\boldsymbol{b}_r$. We consider reduced-order models, $H_r(s)$, that are obtained via projection. That is, we choose full rank matrices $\boldsymbol{V}_r, \boldsymbol{W}_r \in \mathbb{R}^{n\times r}$ such that $\boldsymbol{W}_r^T\boldsymbol{V}_r$ is invertible and define the reduced-order state-space realization with (2) and

$$\boldsymbol{A}_r = (\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{A}\boldsymbol{V}_r,\ \boldsymbol{b}_r = (\boldsymbol{W}_r^T\boldsymbol{V}_r)^{-1}\boldsymbol{W}_r^T\boldsymbol{b},\ \boldsymbol{c}_r = \boldsymbol{V}_r^T\boldsymbol{c}. \tag{3}$$

Within this "projection framework," selection of $\boldsymbol{W}_r$ and $\boldsymbol{V}_r$ completely determines the reduced system – indeed, it is sufficient to specify only the *ranges* of $\boldsymbol{W}_r$ and $\boldsymbol{V}_r$ in order to determine $H_r(s)$. Of particular utility for us is a result by Grimme [6], that gives conditions on $\boldsymbol{W}_r$ and $\boldsymbol{V}_r$ so that the associated reduced-order system, $H_r(s)$, is a *rational Hermite interpolant* to the original system, $H(s)$.

Theorem 1.1 (Grimme [6])**.** *Given $H(s) = \boldsymbol{c}^T(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}$, and r distinct points $s_1, \ldots, s_r \in \mathbb{C}$, let*

$$\boldsymbol{V}_r = [(s_1\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b} \ldots (s_r\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}] \qquad \boldsymbol{W}_r^T = \begin{bmatrix} \boldsymbol{c}^T(s_1\boldsymbol{I}-\boldsymbol{A})^{-1} \\ \vdots \\ \boldsymbol{c}^T(s_r\boldsymbol{I}-\boldsymbol{A})^{-1} \end{bmatrix}. \tag{4}$$

Define the reduced-order model $H_r(s) = \boldsymbol{c}_r^T(s\boldsymbol{I}_r - \boldsymbol{A}_r)^{-1}\boldsymbol{b}_r$ as in (3) Then H_r is a rational Hermite interpolant to H at $s_1, \ldots, s_r$:

$$H(s_i) = H_r(s_i) \qquad \text{and} \qquad H'(s_i) = H_r'(s_i) \quad \text{for} \quad i = 1, \ldots, r. \tag{5}$$

Rational interpolation within this "projection framework" was first proposed by Skelton *et al.* [18],[20],[21]. Later in [6], Grimme established the connection with the rational Krylov method of Ruhe [14].

Significantly, Theorem 1.1 gives an explicit method for computing a reduced-order system that is a Hermite interpolant of the orginal system for nearly *any* set of distinct points, $\{s_1, \ldots, s_r\}$, yet it is not apparent how one should choose these interpolation points in order to assure a high-fidelity reduced-order model in the end. Indeed, the lack of such a strategy had been a major drawback for interpolatory model reduction until recently, when an effective strategy for selecting interpolation points was proposed in [8] yielding reduced-order models that solve

$$\|H - H_r\|_{\mathcal{H}_2} = \min_{dim(\hat{H}_r)=r} \left\|H - \hat{H}_r\right\|_{\mathcal{H}_2}. \tag{6}$$

where the $\mathcal{H}_2$ system norm is defined in the usual way:

$$\|H\|_{\mathcal{H}_2} = \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \mid H(\jmath\omega) \mid^2 d\omega\right)^{1/2}. \tag{7}$$

The optimization problem (6) has been studied extensively, see, for example, [13, 19, 8, 15, 5, 17, 7, 2, 3, 22] and references therein. (6) is a nonconvex optimization problem and finding

global minimizers will be infeasible, typically. Hence, the usual interpretation of (6) involves finding *local* minimizers and a common approach to accomplish this is to construct reduced-order models satisfying first-order necessary optimality conditions. This may be posed either in terms of solutions to Lyapunov equations (e.g., [19, 15, 22]) or in terms of interpolation (e.g., [19, 8, 17, 3]):

Theorem 1.2. *([13, 8]) Given $H(s)$, let $H_r(s)$ be a solution to (6) with simple poles $\hat{\lambda}_1, \ldots, \hat{\lambda}_r$. Then*

$$H(-\hat{\lambda}_i) = H_r(-\hat{\lambda}_i) \quad \text{and} \quad H'(-\hat{\lambda}_i) = H_r'(-\hat{\lambda}_i) \quad \text{for} \quad i = 1, \ldots, r. \tag{8}$$

That is, any $\mathcal{H}_2$-optimal reduced order model of order r with simple poles will be a Hermite interpolant to $H(s)$ at the reflected image of the reduced poles through the origin.

Although this result might appear to reduce the problem of $\mathcal{H}_2$-optimal model approximation to a straightforward application of Theorem 1.1 to calculate a Hermite interpolant on the set of reflected poles, $\{-\hat{\lambda}_1, \ldots, -\hat{\lambda}_r\}$, these pole locations will not be known *a priori*. Nonetheless, these pole locations can be determined efficiently with the Iterative Rational Krylov Algorithm (IRKA) of Gugercin *et al.* [8]. Starting from an arbitrary initial selection of interpolation points, IRKA iteratively corrects the interpolation points until (8) is satisfied. A brief sketch of IRKA is given below.

Algorithm IRKA. *Iterative Rational Krylov Algorithm [8]*

Given a full-order $H(s)$, a reduction order r, and convergence tolerance tol.

1. Make an initial selection of r distinct interpolation points, $\{s_i\}_1^r$, that is closed under complex conjugation.
2. Construct $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$ as in (4).
3. while (relative change in $\{s_i\} > tol$)
   a.) $\boldsymbol{A}_r = (\boldsymbol{W}_r^T \boldsymbol{V}_r)^{-1} \boldsymbol{W}_r^T \boldsymbol{A} \boldsymbol{V}_r$
   b.) Solve $r \times r$ eigenvalue problem $\boldsymbol{A}_r \mathbf{u} = \lambda \mathbf{u}$ and assign $s_i \leftarrow -\lambda_i(\boldsymbol{A}_r)$ for $i = 1, \ldots, r$.
   c.) Update $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$ with new s_i's using (4).
4. $\boldsymbol{A}_r = (\boldsymbol{W}_r^T \boldsymbol{V}_r)^{-1} \boldsymbol{W}_r^T \boldsymbol{A} \boldsymbol{V}_r$, $\quad \boldsymbol{b}_r = (\boldsymbol{W}_r^T \boldsymbol{V}_r)^{-1} \boldsymbol{W}_r^T \boldsymbol{b}$, and $\boldsymbol{c}_r = \boldsymbol{V}_r^T \boldsymbol{c}$.

IRKA has been remarkably successful in producing high fidelity reduced-order approximations and has been successfully applied to finding $\mathcal{H}_2$-optimal reduced models for systems of high order, $n > 160,000$, see [9]. For details on IRKA, see [8].

Notwithstanding typically observed rapid convergence of the IRKA iteration to interpolation points that generally yield high quality reduced models, no convergence theory for IRKA has yet been established. Evidently from the description above, IRKA may be viewed as a fixed point iteration with fixed points coinciding with the stationary points of the $\mathcal{H}_2$ minimization problem. Saddle points and local maxima of the $\mathcal{H}_2$ minimization problem are known to be repellent [11]. However, despite effective performance in practice, it has not yet been established that local minima are attractive fixed points.

In this paper, we give a proof of this for the special case of state-space-symmetric systems and establish the convergence of IRKA for this class of systems.

2 State-Space-Symmetric Systems

Definition 1. $H(s)=\boldsymbol{c}^T(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}$ is *state-space-symmetric* (SSS) if $\boldsymbol{A}=\boldsymbol{A}^T$ and $\boldsymbol{c}=\boldsymbol{b}$.

SSS systems appear in many important applications such as in the analysis of RC circuits and in inverse problems involving 3D Maxwell's equations [4].

A closely related class of systems is the class of zero-interlacing-pole (ZIP) systems.

Definition 2. A system $H(s) = K\dfrac{\prod_{i=1}^{n-1}(s-z_i)}{\prod_{j=1}^{n}(s-\lambda_j)}$ is a *strictly proper ZIP system* provided that

$$0 > \lambda_1 > z_1 > \lambda_2 > z_2 > \lambda_3 > \cdots > z_{n-1} > \lambda_n.$$

The following relation serves to characterize ZIP systems.

Proposition 2.1. *[16] $H(s)$ is a strictly proper ZIP system if and only if $H(s)$ can be written as $H(s)=\displaystyle\sum_{i=1}^{n}\frac{b_i}{s-\lambda_i}$ with $\lambda_i<0$, $b_i>0$, and $\lambda_i\neq\lambda_j$ for all $i\neq j$.*

The next result clarifies the relationship between SSS and ZIP systems.

Lemma 2.1. *[12] Let $H(s)$ be SSS. Then $H(s)$ is minimal if and only if the poles of $H(s)$ are distinct. Moreover, every SSS system has a SSS minimal realization with distinct poles, and is therefore a strictly proper ZIP system.*

It can easily be verified from the implementation of IRKA given above, that for SSS systems, the relationship $\boldsymbol{V}_r=\boldsymbol{W}_r$ is maintained throughout the iteration, and the final reduced-order model at Step 4 of IRKA can be obtained by

$$\boldsymbol{A}_r = \boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r \quad \boldsymbol{b}_r = \boldsymbol{c}_r = \boldsymbol{Q}_r^T\boldsymbol{b}, \tag{9}$$

where $\boldsymbol{Q}_r$ is an orthonormal basis for $\boldsymbol{V}_r$; the reduced system resulting from IRKA is also SSS.

3 The Main Result

Theorem 3.1. *Let IRKA be applied to a minimal SSS system $H(s)$. Then every fixed point of IRKA which is a local minimizer is locally attractive. In other words, IRKA is a locally convergent fixed point iteration to a local minimizer of the $\mathcal{H}_2$ optimization problem.*

To proceed with the proof of Theorem 3.1, we need four intermediate lemmas. The first lemma provides insight into the structure of the zeros of the error system resulting from reducing a SSS system.

Lemma 3.1. *Let $H(s)$ be a SSS system of order n. If $H_r(s)$ is a ZIP system that interpolates $H(s)$ at $2r$ points $s_1, s_2, \dots, s_{2r}$, not necessarily distinct, in $(0,\infty)$, then all the remaining zeros of the error system lie in $(-\infty, 0)$.*

Proof. By Lemma 2.1, we may assume that $H(s)$ is a strictly proper ZIP systems. Since $H(s)$ is a strictly proper ZIP system, its poles are simple and all its residues are positive. Let $\lambda_i < 0, \phi_i > 0$, for $i = 1, \dots, n$ be the poles and residues of $H(s)$, respectively. Now let

$$R(s) = \prod_{i=1}^{2r}(s - s_i), \quad P(s) = \prod_{i=1}^{n-r-1}(s + z_i), \quad Q(s) = \prod_{i=1}^{n}(s - \lambda_i), \quad \tilde{Q}(s) = \prod_{i=1}^{r}(s - \tilde{\lambda}_i),$$

where $\tilde{\lambda}_i$, s_i, and z_i are, respectively, the poles of $H_r(s)$, the interpolation points, and the remaining zeros of the error system. Then for some constant K, $H(s) - H_r(s) = K\frac{P(s)R(s)}{Q(s)\tilde{Q}(s)}$. First suppose that $\{\lambda_i\}_{i=1}^{n} \cap \{\tilde{\lambda}_k\}_{k=1}^{r} = \emptyset$. Then for each λ_j, $j = 1, \dots, n$,

$$\operatorname{Res}(H(s) - H_r(s); \lambda_j) = K\frac{P(\lambda_j)R(\lambda_j)}{\prod\limits_{\substack{i=1 \\ \lambda_i \neq \lambda_j}}^{n}(\lambda_j - \lambda_i)\tilde{Q}(\lambda_j)} = \phi_i > 0. \tag{10}$$

Thus, $\operatorname{sgn}(KP(\lambda_j)) = (-1)^{j-1}\operatorname{sgn}(\tilde{Q}(\lambda_j))$ where $\operatorname{sgn}(\alpha)$ denotes the sign of α. Now if $(-1)^{j-1}\operatorname{sgn}(\tilde{Q}(\lambda_j)) = (-1)^j(\operatorname{sgn}(\tilde{Q}(\lambda_{j+1}))$, then $-\operatorname{sgn}(\tilde{Q}(\lambda_j)) = \operatorname{sgn}(\tilde{Q}(\lambda_{j+1}))$, so $\tilde{Q}(s)$ must change sign on the interval $[\lambda_{j+1}, \lambda_j]$. Since $\tilde{Q}(s)$ is a polynomial of degree r, and $r < n$, $\tilde{Q}(s)$ can switch signs at most r times, else $\tilde{Q}(s) \equiv 0$. But this means there are at least $n - r - 1$ intervals $[\lambda_{j_k+1}, \lambda_{j_k}]$, for $k = 1, \dots, n - r - 1$, for which $\operatorname{sgn}(\tilde{Q}(\lambda_{j_k})) = \operatorname{sgn}(\tilde{Q}(\lambda_{j_k+1}))$, and therefore $\operatorname{sgn}(KP(\lambda_{j_k})) = -\operatorname{sgn}(KP(\lambda_{j_k+1}))$. So $KP(s)$ must change sign over at least $n - r - 1$ intervals, and therefore has at least $n - r - 1$ zeros on $[\lambda_n, \lambda_1]$. Again, since the error is not identically zero when $r < n$, and the degree of $KP(s)$ is $n - r - 1$, this implies that all the zeros of $KP(s)$ lie in $(-\infty, 0)$.

Suppose with some $p \leq r$, $\lambda_{i_j} = \tilde{\lambda}_{k_j}$ for $j = 1, \dots, p$. Observe from partial fraction expansions of $H(s)$ and $H_r(s)$ that the error can be written as a rational function of degree $n + r - p - 1$

over degree $n+r-p$ with distinct poles. n of these poles belong to $H(s)$ and the remaining $r-p$ come from the poles of $H_r(s)$ that are distinct from the poles of $H(s)$. Now let

$$R(s) = \prod_{i=1}^{2r}(s-s_i), \quad P(s) = \prod_{i=1}^{n-r-p-1}(s+z_i), \quad Q(s) = \prod_{i=1}^{n}(s-\lambda_i), \quad \tilde{Q}(s) = \prod_{l=1}^{r-p}(s-\tilde{\lambda}_{k_l}),$$

where $\{\tilde{\lambda}_{k_l}\}_{l=1}^{r-p} = \{\tilde{\lambda}_k\}_{k=1}^{r} \backslash \{\lambda_i\}_{i=1}^{n}$. Hence, $H(s)-H_r(s) = K\frac{P(s)R(s)}{Q(s)\tilde{Q}(s)}$. Observe that there are at most $2p$ subintervals of the form $[\lambda_{i^*}, \lambda_{i^*+1}]$ or $[\lambda_{i^*-1}, \lambda_{i^*}]$, where $\lambda_{i^*} \in \{\lambda_i\}_{i=1}^{n} \cap \{\tilde{\lambda}_k\}_{k=1}^{r}$. It follows that there are at least $n-2p-1$ subintervals of the form $[\lambda_i, \lambda_{i+1}]$, where $\lambda_i, \lambda_{i+1} \notin \{\lambda_i\}_{i=1}^{n} \cap \{\tilde{\lambda}_k\}_{k=1}^{r}$. On each such subinterval for which this is the case, we have

$$\operatorname{Res}(H(s)-H_r(s);\lambda_i) = K\frac{P(\lambda_i)R(\lambda_i)}{\prod\limits_{\substack{j=1\\ \lambda_j\neq\lambda_i}}^{n}(\lambda_i-\lambda_j)\tilde{Q}(\lambda_i)} = \phi_i > 0. \tag{11}$$

So $\operatorname{sgn}(KP(\lambda_i)) = (-1)^{i-1}\operatorname{sgn}(\tilde{Q}(\lambda_i))$. By the same argument as above where the poles of $H(s)$ and $H_r(s)$ are distinct, either $\tilde{Q}(s)$ or $P(s)$ has a zero on the interval $[\lambda_i, \lambda_{i+1}]$. Since $\tilde{Q}(s)$ has at most $r-p$ zeros, this means that there are at least $n-2p-1-(r-p) = n-p-r-1$ subintervals between poles of $H(s)$ where $P(s)$ has zeros. Hence, the lemma is proved. ☐

Lemma 3.2. *Let $H(s) = \boldsymbol{b}^T(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}$ be SSS, and $H_r(s) = \boldsymbol{b}_r^T(s\boldsymbol{I}_r-\boldsymbol{A}_r)^{-1}\boldsymbol{b}_r$ be any reduced order model of $H(s)$ constructed by a compression of $H(s)$, i.e., $\boldsymbol{A}_r = \boldsymbol{Q}_r^T\boldsymbol{A}_r\boldsymbol{Q}_r$, $\boldsymbol{b}_r = \boldsymbol{Q}_r^T\boldsymbol{b}$. Then for any $s \geq 0$, $H(s)-H_r(s) \geq 0$.*

Proof. Pick any $s \geq 0$. Then $(s\boldsymbol{I}_n-\boldsymbol{A})$ is symmetric, positive definite and has a Cholesky decomposition, $(s\boldsymbol{I}_n-\boldsymbol{A}) = \boldsymbol{L}\boldsymbol{L}^T$. Define $\boldsymbol{Z}_r = \boldsymbol{L}^T\boldsymbol{Q}_r$. Then

$$\begin{aligned} H(s)-H_r(s) &= \boldsymbol{b}^T\left[(s\boldsymbol{I}_n-\boldsymbol{A})^{-1} - \boldsymbol{Q}_r\left(\boldsymbol{Q}_r^T(s\boldsymbol{I}_n-\boldsymbol{A})\boldsymbol{Q}_r\right)^{-1}\boldsymbol{Q}_r^T\right]\boldsymbol{b} \\ &= (\boldsymbol{L}^{-1}\boldsymbol{b})^T\left[\boldsymbol{I}-\boldsymbol{Z}_r\left(\boldsymbol{Z}_r^T\boldsymbol{Z}_r\right)^{-1}\boldsymbol{Z}_r^T\right](\boldsymbol{L}^{-1}\boldsymbol{b}). \end{aligned}$$

Note the last bracketed expression is an orthogonal projector onto $\mathsf{Ran}(\boldsymbol{Z}_r)^{\perp}$, hence is positive semidefinite and the conclusion follows. ☐

Our convergence analysis of IRKA will use its formulation as a fixed-point iteration. The analysis will build on the framework of [11]. Let

$$H(s) = \sum_{i=1}^{n}\frac{\phi_i}{s-\lambda_i} \quad \text{and} \quad H_r(s) = \sum_{j=1}^{r}\frac{\tilde{\phi}_j}{s-\tilde{\lambda}_j} \tag{12}$$

be the partial fraction decompositions of $H(s)$, and $H_r(s)$, respectively. Given a set of r interpolation points $\{s_i\}_{i=1}^{r}$, identify the set with a vector $\boldsymbol{s} = [s_1,\ldots,s_r]^T$. Construct an

interpolatory reduced order model $H_r(s)$ from $\boldsymbol{s}$ as in Theorem 1.1 and identify $\{\tilde{\lambda}_i\}_{i=1}^r$ with a vector $\tilde{\boldsymbol{\lambda}} = [\tilde{\lambda}_1, \ldots, \tilde{\lambda}_r]^T$. Then define the function $\lambda : \mathbb{C}^r \to \mathbb{C}^r$ by $\lambda(\boldsymbol{s}) = -\tilde{\boldsymbol{\lambda}}$. Aside from ordering issues, this function is well defined, and the IRKA iteration converges when $\lambda(\boldsymbol{s}) = \boldsymbol{s}$. Thus convergence of IRKA is equivalent to convergence of a fixed point iteration on the function $\lambda(\boldsymbol{s})$. Similar to $\boldsymbol{s}$ and $\tilde{\boldsymbol{\lambda}}$, let $\tilde{\boldsymbol{\phi}} = [\tilde{\phi}_1, \ldots, \tilde{\phi}_r]^T$. Having identified $H_r(s)$ with its poles and residues, the optimal $\mathcal{H}_2$ model reduction problem may be formulated in terms of minimizing the cost function $\boldsymbol{\mathcal{J}}(\tilde{\boldsymbol{\phi}}, \lambda(\boldsymbol{s})) = \|H - H_r\|_{\mathcal{H}_2}^2$, where

$$\boldsymbol{\mathcal{J}}(\tilde{\boldsymbol{\phi}}, \lambda(\boldsymbol{s})) = \sum_{i=1}^{n} \phi_i (H(\lambda_i) - H_r(\lambda_i)) + \sum_{j=1}^{r} \tilde{\phi}_j (H(\tilde{\lambda}_j) - H_r(\tilde{\lambda}_j)) \tag{13}$$

See [8] for a derivation of (13). Define the matrices $\boldsymbol{S}_{11}, \boldsymbol{S}_{12}, \boldsymbol{S}_{22} \in \mathbb{R}^{r\times r}$ as

$$[\boldsymbol{S}_{11}]_{i,j} = -(\tilde{\lambda}_i + \tilde{\lambda}_j)^{-1}, \quad [\boldsymbol{S}_{12}]_{i,j} = -(\tilde{\lambda}_i + \tilde{\lambda}_j)^{-2} \quad \text{and} \quad [\boldsymbol{S}_{22}]_{i,j} = -2(\tilde{\lambda}_i + \tilde{\lambda}_j)^{-3}$$

for $i, j = 1, \ldots, r$. Also, define $\boldsymbol{R}, \boldsymbol{E} \in \mathbb{R}^{r\times r}$:

$$\boldsymbol{R} = \operatorname{diag}(\{\tilde{\phi}_1, \ldots, \tilde{\phi}_r\}), \quad \text{and} \quad \boldsymbol{E} = \operatorname{diag}(\{H''(-\tilde{\lambda}_1) - H_r''(-\tilde{\lambda}_1), \ldots, H''(-\tilde{\lambda}_r) - H_r''(-\tilde{\lambda}_r)\}.$$

Lemma 3.3. *Let $H(s)$ be SSS and let $H_r(s)$ be an IRKA interpolant. Then $\boldsymbol{E}$ is positive definite at any fixed point of $\lambda(\boldsymbol{s})$.*

Proof. By Lemma 3.2, $H(s) - H_r(s) \geq 0$ for all $s \in [0, \infty)$. Thus the points $H(-\tilde{\lambda}_i) - H_r(-\tilde{\lambda}_i)$ are local minima of $H(s) - H_r(s)$ on $[0, \infty)$ for $i = 1, \ldots, r$. It then follows that $H''(-\tilde{\lambda}_i) - H_r''(-\tilde{\lambda}_i) \geq 0$. But by Lemma 3.1, $H(s) - H_r(s)$ has exactly $2r$ zeros in $\mathbb{C}_+$, so $H''(-\tilde{\lambda}_i) - H_r''(-\tilde{\lambda}_i) > 0$ for $i = 1, \ldots, r$. ∎

Lemma 3.4. *The matrix $\tilde{\boldsymbol{S}} = \begin{bmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{12} & \boldsymbol{S}_{22} \end{bmatrix}$ is positive definite.*

Proof. We will show that for any non-zero vector $\boldsymbol{z} = [z_1, z_2, \ldots, z_{2r}]^T \in \mathbb{R}^{2r}$

$$\boldsymbol{z}^T \boldsymbol{S} \boldsymbol{z} = \int_0^\infty \left[\sum_{i=1}^{r} z_i e^{\tilde{\lambda}_i t} - t\Big(\sum_{i=1}^{r} z_{r+i} e^{\tilde{\lambda}_i t} \Big) \right]^2 \mathrm{d}t > 0.$$

Define $\boldsymbol{z}_r = [z_1, z_2, \ldots, z_r]^T \in \mathbb{R}^r$ and $\boldsymbol{z}_{2r} = [z_{r+1}, z_{r+2}, \ldots, z_{2r}]^T \in \mathbb{R}^r$. Then

$$\boldsymbol{z}^T \boldsymbol{S} \boldsymbol{z} = \boldsymbol{z}_r^T \boldsymbol{S}_{11} \boldsymbol{z}_r + 2\boldsymbol{z}_r^T \boldsymbol{S}_{12} \boldsymbol{z}_{2r} + \boldsymbol{z}_{2r}^T \boldsymbol{S}_{22} \boldsymbol{z}_{2r} \tag{14}$$

Let $\boldsymbol{\Lambda} = \operatorname{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_r)$ and $\boldsymbol{u}$ be a vector of r ones. Note that $\boldsymbol{S}_{11}$ solves the Lyapunov equation $\boldsymbol{\Lambda} \boldsymbol{S}_{11} + \boldsymbol{S}_{11} \boldsymbol{\Lambda} + \boldsymbol{u}\boldsymbol{u}^T = \boldsymbol{0}$. Thus,

$$\boldsymbol{z}_r^T \boldsymbol{S}_{11} \boldsymbol{z}_r = \int_0^\infty \boldsymbol{z}_r^T e^{\boldsymbol{\Lambda} t} \boldsymbol{u}\boldsymbol{u}^T e^{\boldsymbol{\Lambda} t} \boldsymbol{z}_r \, \mathrm{d}t = \int_0^\infty \Big(\sum_{i=1}^{r} z_i e^{\tilde{\lambda}_i t} \Big)^2 \mathrm{d}t \tag{15}$$

Similarly, $\boldsymbol{S}_{12}$ solves $\boldsymbol{\Lambda}\boldsymbol{S}_{12} + \boldsymbol{S}_{12}\boldsymbol{\Lambda} - \boldsymbol{S}_{11} = \boldsymbol{0}$. An application of integration by parts gives:

$$\begin{aligned}
\int_0^\infty t\Big(\sum_{i=1}^{r} z_i e^{\tilde{\lambda}_i t}\Big)\Big(\sum_{i=1}^{r} z_{r+i} e^{\tilde{\lambda}_i t}\Big)\,\mathrm{d}t &= \int_0^\infty t(\boldsymbol{z}_r^T (e^{\boldsymbol{\Lambda}t}\boldsymbol{u}\boldsymbol{u}^T e^{\boldsymbol{\Lambda}t})\boldsymbol{z}_{2r})\,\mathrm{d}t \\
&= \boldsymbol{z}_r^T\Big[-t(e^{\boldsymbol{\Lambda}t}\boldsymbol{S}_{11}e^{\boldsymbol{\Lambda}t})\Big]_0^\infty \boldsymbol{z}_{2r} + \boldsymbol{z}_r^T\Big(\int_0^\infty e^{\boldsymbol{\Lambda}t}\boldsymbol{S}_{11}e^{\boldsymbol{\Lambda}t}\,\mathrm{d}t\Big)\boldsymbol{z}_{2r} \\
&= -\boldsymbol{z}_r^T\boldsymbol{S}_{12}\boldsymbol{z}_{2r} \qquad (16)
\end{aligned}$$

Finally, note that $\boldsymbol{S}_{22}$ solves $\boldsymbol{\Lambda}\boldsymbol{S}_{22} + \boldsymbol{S}_{22}\boldsymbol{\Lambda} - 2\boldsymbol{S}_{12} = \boldsymbol{0}$. Repeated applications of integration by parts then yields the equality:

$$\boldsymbol{z}_{2r}^T\boldsymbol{S}_{22}\boldsymbol{z}_{2r} = \int_0^\infty t^2\Big(\sum_{i=1}^{r} z_{r+1}e^{\tilde{\lambda}_i t}\Big)^2\,\mathrm{d}t \qquad (17)$$

Combining equations (14), (15), (16), and (17) gives the desired results since

$$\boldsymbol{z}^T\boldsymbol{S}\boldsymbol{z} = \int_0^\infty \Big[\sum_{i=1}^{r} z_i e^{\tilde{\lambda}_i t} - t\Big(\sum_{i=1}^{r} z_{r+i}e^{\tilde{\lambda}_i t}\Big)\Big]^2\,\mathrm{d}t.$$

□

Then it follows that the Schur complement $\boldsymbol{S}_{22} - \boldsymbol{S}_{12}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}$ of $\tilde{\boldsymbol{S}}$ is also positive definite. With the setup above, we may now commence with the proof of Theorem 3.1.

Proof of Theorem 3.1: It suffices to show that for any fixed point which is a local minimizer of $\boldsymbol{\mathcal{J}}(\tilde{\boldsymbol{\phi}}, \lambda(\boldsymbol{s}))$, the eigenvalues of the Jacobian of $\lambda(\boldsymbol{s})$ are bounded in magnitude by 1. As shown in [11], the Jacobian of $\lambda(\boldsymbol{s})$ can be written as $-\boldsymbol{S}_c^{-1}\boldsymbol{K}$ where

$$\boldsymbol{S}_c = \boldsymbol{S}_{22} - \boldsymbol{S}_{12}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12} \qquad \text{and} \qquad \boldsymbol{K} = \boldsymbol{E}\boldsymbol{R}^{-1}.$$

First off, note that from Lemma 3.3, and the fact that $H(s)$ is a ZIP system by Lemma 2.1, $\boldsymbol{K}$ is positive definite. Evaluating the pencil $\boldsymbol{K} - \lambda\boldsymbol{S}_c$ at $\lambda = 1$ gives

$$\boldsymbol{\Phi} = -\boldsymbol{S}_{22} + \boldsymbol{E}\boldsymbol{R}^{-1} + \boldsymbol{S}_{12}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12},$$

This pencil is regular since $\boldsymbol{S}_c$ is positive definite by Lemma 3.4, and therefore $\det(\boldsymbol{K} - \lambda\boldsymbol{S}_c)$ is zero if and only if $\det(\boldsymbol{S}_c^{-1}\boldsymbol{K} - \lambda\boldsymbol{I}) = 0$.

Let $\nabla^2\boldsymbol{\mathcal{J}}$ denote the Hessian of the cost function $\boldsymbol{\mathcal{J}}(\tilde{\boldsymbol{\phi}}, \lambda(\boldsymbol{s}))$. As shown in [11], $\nabla^2\boldsymbol{\mathcal{J}}$ can be written as

$$\nabla^2\boldsymbol{\mathcal{J}} = \begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}\end{bmatrix}\boldsymbol{M}\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}\end{bmatrix}, \quad \text{where} \quad \boldsymbol{M} = \begin{bmatrix}\boldsymbol{S}_{11} & \boldsymbol{S}_{12}\\ \boldsymbol{S}_{12} & \boldsymbol{S}_{22} - \boldsymbol{E}\boldsymbol{R}^{-1}\end{bmatrix}.$$

Note that $-\boldsymbol{\Phi}$ is the Schur complement of $\boldsymbol{M}$. Hence, if the fixed point is a local minimimum, then $-\boldsymbol{\Phi}$ must be positive definite and so for $\lambda = 1$ the pencil is negative definite. Since both $\boldsymbol{K}$ and $\boldsymbol{S}_c$ are positive definite, there exists a nonsingular transformation $\boldsymbol{Z}$ by which the quadratic form $\boldsymbol{y}^T(\boldsymbol{K} - \lambda \boldsymbol{S}_c)\boldsymbol{y}$ is transformed into $\boldsymbol{z}^T(\boldsymbol{\Lambda} - \lambda \boldsymbol{I})\boldsymbol{z}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix formed from the solutions of

$$\det(\boldsymbol{K} - \lambda \boldsymbol{S}_c) = 0. \tag{18}$$

Thus, the solutions of (18) correspond to the eigenvalues of $\boldsymbol{S}_c^{-1}\boldsymbol{K}$. $\boldsymbol{\Lambda} - \boldsymbol{I}$ must be negative definite since $\boldsymbol{\Phi}$ is, and therefore the eigenvalues of the $\boldsymbol{S}_c^{-1}\boldsymbol{K}$ must be real-valued and less than one. Furthermore, note that $\boldsymbol{P} = \boldsymbol{S}_c^{-1}\boldsymbol{K}$ solves the Lyapunov equation

$$\boldsymbol{P}\boldsymbol{S}_c^{-1} + \boldsymbol{S}_c^{-1}\boldsymbol{P}^T = 2\boldsymbol{S}_c^{-1}\boldsymbol{K}\boldsymbol{S}_c^{-1},$$

so by the standard inertia result, all the eigenvalues of $\boldsymbol{S}_c^{-1}\boldsymbol{K}$ are positive, and the desired result follows. $\square$

References

[1] A.C. Antoulas. *Approximation of Large-Scale Dynamical Systems (Advances in Design and Control)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005.

[2] C.A. Beattie and S. Gugercin. Krylov-based minimization for optimal $\mathcal{H}_2$ model reduction. *46th IEEE Conference on Decision and Control*, pages 4385–4390, Dec. 2007.

[3] C.A. Beattie and S. Gugercin. A trust region method for optimal $\mathcal{H}_2$ model reduction. *48th IEEE Conference on Decision and Control*, Dec. 2009.

[4] V. Druskin, L. Knizhnerman, and M. Zaslavsky. Solution of large scale evolutionary problems using rational Krylov subspaces with optimized shifts. *SIAM Journal on Scientific Computing*, 31(5):3760–3780, 2009.

[5] P. Fulcheri and M. Olivi. Matrix rational $\mathcal{H}_2$ approximation: a gradient algorithm based on Schur analysis. *SIAM Journal on Control and Optimization*, 36(6):2103–2127, 1998.

[6] E.J. Grimme. *Krylov projection methods for model reduction.* PhD thesis, University of Illinois, 1997.

[7] S. Gugercin. An iterative rational Krylov algorithm (IRKA) for optimal $\mathcal{H}_2$ model reduction. In *Householder Symposium XVI*, Seven Springs Mountain Resort, PA, USA, May 2005.

[8] S. Gugercin, A.C. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30(2):609–638, 2008.

[9] A.R. Kellems, D. Roos, N. Xiao, and S.J. Cox. Low-dimensional, morphologically accurate models of subthreshold membrane potential. *Journal of Computational Neuroscience*, 27(2):161–176, 2009.

[10] J.G. Korvink and E.B. Rudnyi. Oberwolfach benchmark collection. In P. Benner, V. Mehrmann, and D. C. Sorensen, editors, *Dimension Reduction of Large-Scale Systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 311–315. Springer-Verlag, Berlin/Heidelberg, Germany, 2005.

[11] W. Krajewski, A. Lepschy, M. Redivo-Zaglia, and U. Viaro. A program for solving the $\mathcal{L}_2$ reduced-order model problem with fixed denominator degree. *Numerical Algorithms*, 9(2):355–377, 1995.

[12] W.Q. Liu, V. Sreeram, and K.L. Teo. Model Reduction for State-space Symmetric Systems. *Systems & Control Letters*, 34:209–215, 1998.

[13] L. Meier III and D. Luenberger. Approximation of linear constant systems. *Automatic Control, IEEE Transactions on*, 12(5):585–588, 1967.

[14] A. Ruhe. Rational Krylov: A practical algorithm for large sparse nonsymmetric matrix pencils. *SIAM Journal on Scientific Computing*, 19(5):1535–1551, 1998.

[15] J.T. Spanos, M.H. Milman, and D.L. Mingori. A new algorithm for L^2 optimal model reduction. *Automatica*, 28(5):897–909, 1992.

[16] B. Srinivasan and P. Myszkorowski. Model reduction of systems with zeros interlacing the poles. *Systems & Control Letters*, 30(1):19–24, 1997.

[17] P. Van Dooren, K.A. Gallivan, and P.A. Absil. $\mathcal{H}_2$-optimal model reduction of MIMO systems. *Applied Mathematics Letters*, 21(12):1267–1273, 2008.

[18] C.D. Villemagne and R.E. Skelton. Model reduction using a projection formulation. *Intl. J. Contr*, 46:2141–2169, 1987.

[19] D.A. Wilson. Optimum solution of model-reduction problem. *Proc. IEE*, 117(6):1161–1165, 1970.

[20] A. Yousuff and R.E. Skelton. Covariance Equivalent Realizations with Application to Model Reduction of Large Scale Systems. *Control and Dynamic Systems*, 22, 1984.

[21] A. Yousuff, D.A. Wagie, and R.E. Skelton. Linear system approximation via covariance equivalent realizations. *Journal of mathematical analysis and applications*, 106(1):91–115, 1985.

[22] D. Zigic, L.T. Watson, and C.A. Beattie. Contragredient transformations applied to the optimal projection equations. *Linear algebra and its applications*, 188:665–676, 1993.